ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Statistical analysis of the LIUQE algorithm

Lab Report

Date	:	January 10, 2015
Course	:	TP IV, CRPP
Name	:	Nathan Dupertuis
Topic	:	LIUQE algorithm
Professor	:	J.-M. Moret

Contents

Abstract

The new LIUQE algorithm computes a real-time reconstitution of the plasma in TCV, which gives interesting prospectives of tokamak control. Its computation relies on multiple linear regression. The study is introduced with theoretical features on regression and LIUQE algorithm. Covariance matrix and standard error values for parameters of regression are then computed on a TCV shot. The analysis of results focuses on the quality of the regression model. A two basis function model appears finally to be the most relevant parametrization of the algorithm.

1 Introduction

1.1 Elements of notation for regression analysis

1.1.1 Minimizing the residual and variance ellipsoids

The residual of a multiple linear regression can be computed in matrix notation:

$$RSS(\boldsymbol{\beta}) = \sum_{i=1}^{n}(y_i - \boldsymbol{x_i}^T\boldsymbol{\beta})^2 = (\boldsymbol{Y}-\boldsymbol{X\beta})^T(\boldsymbol{Y}-\boldsymbol{X\beta}) \tag{1}$$

For simplicity the residual will be denoted by $\chi^2 := RSS(\boldsymbol{\beta})$.
The estimate of the vector of regression coefficients is expressed by:

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{Y}) \tag{2}$$

This estimate is the vector that minimizes the residual. Another useful expression is easily derivable from the last two result, isolating the minimal value of χ^2:

$$\chi^2 = \chi^2_{min} + (\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^T(\boldsymbol{X}^T\boldsymbol{X})(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}) \tag{3}$$

with

$$\chi^2_{min} = (\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})^T(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}) \tag{4}$$

The second term on the right of equation 3 correspond to the parametrization of an ellipsoid centered on the point of coordinates $\hat{\boldsymbol{\beta}}$, in the space of coefficients $\{\beta_i\}$ of dimension p. In order to exhibit the equation of an ellipsoid, a *singular value decomposition* can be applied, leading to a diagonalisation of the $\boldsymbol{X}^T\boldsymbol{X}$ matrix:

$$\chi^2 = \chi^2_{min} + (\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^T(\boldsymbol{U}^T\boldsymbol{S}\boldsymbol{U})(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}) \tag{5}$$

$\boldsymbol{S}$ is the diagonal matrix and $\boldsymbol{U}$ is a unitary matrix. This equation can be understood with the following: the columns of $\boldsymbol{U}$ give the axes of the ellipsoid in the coefficient space. Those vectors (noted $\boldsymbol{U}_i$ for the i-th column) constitutes an orthonormal basis as $\boldsymbol{U}$ is unitary. The eigenvalues on the diagonal of $\boldsymbol{S}$ give the length of the semi-axes of the ellipsoid.

1.1.2 Variance and correlations

In matrix notation, the covariance matrix for the $\hat{\beta}_i$ parameters can be expressed as follows:

$$\boldsymbol{covar}(\hat{\boldsymbol{\beta}}) = \sigma^2(\boldsymbol{X}^T\boldsymbol{X})^{-1}$$

and its estimation:

$$\widehat{\boldsymbol{covar}}(\hat{\boldsymbol{\beta}}) = \hat{\sigma}^2(\boldsymbol{X}^T\boldsymbol{X})^{-1} \tag{6}$$

with the global estimator of variance (where p is the number of parameters in the model):

$$\hat{\sigma}^2 = \frac{\chi^2_{min}}{n-p} \tag{7}$$

This covariance matrix is normalizable in order to show the correlations between parameters, avoiding the problem of different dimensions and order of magnitude. The normalization consists into dividing a correlation term between two parameters by the corresponding diagonal variances:

$$\left(\widehat{\boldsymbol{N}}(\hat{\boldsymbol{\beta}})\right)_{ij} := \frac{\left(\widehat{\boldsymbol{covar}}(\hat{\boldsymbol{\beta}})\right)_{ij}}{\sqrt{\left(\widehat{\boldsymbol{covar}}(\hat{\boldsymbol{\beta}})\right)_{ii} \cdot \left(\widehat{\boldsymbol{covar}}(\hat{\boldsymbol{\beta}})\right)_{jj}}} \tag{8}$$

1.1.3 Relationship between the covariance matrix and the ellipsoids

It is possible to show that the covariance matrix diagonal gives the exact same maximal values of standard error as the projection of the standard error ellipsoid on the parameters axes, in the case $\chi = 2\chi^2$.
Noting that the points of the ellipsoid surface with maximal standard errors with regard to a parameter axis are those where the gradient is parallel to the parameter axis, one obtains:

$$\nabla\chi^2 = 2h_i\widehat{\boldsymbol{e}}_i \tag{9}$$

where $\widehat{\boldsymbol{e}}_i$ is the unit vector of the parameter axis and $2h_i$ is an arbitrary coefficient with convenient notation.
From equation 3 in the case $\chi = 2\chi^2$, the following comes:

$$\nabla\chi^2 = 2(\boldsymbol{X}^T\boldsymbol{X})(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}) \tag{10}$$

i.e. the vector $\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}$ is known:

$$\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}} = h_i(\boldsymbol{X}^T\boldsymbol{X})^{-1}\widehat{\boldsymbol{e}}_i \tag{11}$$

This vectors goes from the center of the ellipsoid at $\widehat{\boldsymbol{\beta}}$ to the point $\boldsymbol{\beta}$ where the gradient is null. Replacing it in equation 3, one gets:

$$\chi^2 = 2\chi^2_{min} = \chi^2_{min} + h_i^2\widehat{\boldsymbol{e}}_i^T(\boldsymbol{X}^T\boldsymbol{X})^{-1}\widehat{\boldsymbol{e}}_i$$

or, after the writing of the products with $\widehat{\boldsymbol{e}}_i$ vectors:

$$\chi^2_{min} = h_i^2\sum_{kl}\delta_{ik}(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{kl}\delta_{li} = h_i^2(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{ii}$$

So the coefficient has now the expression:

$$h_i = \frac{\chi_{min}}{\sqrt{(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{ii}}} \tag{12}$$

The refreshment of equation 11 gives:

$$\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}} = \frac{\chi_{min}}{\sqrt{(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{ii}}}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\widehat{\boldsymbol{e}}_i \tag{13}$$

In order to get the maximal standard error for the i-th parameter, one has to compute the corresponding component of equation 13:

$$\begin{aligned}
\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i &= \frac{\chi_{min}}{\sqrt{(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{ii}}}\sum_k(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{ik}\delta_{ki} \\
&= \frac{\chi_{min}}{\sqrt{(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{ii}}}(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{ii} \\
&= \sqrt{(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{ii}}\chi_{min} \\
&= \sqrt{\widehat{\boldsymbol{covar}}(\hat{\boldsymbol{\beta}})_{ii}}
\end{aligned}$$

Consequently, in the case $\chi = 2\chi^2$, the covariance matrix diagonal gives the exact same maximal values of standard error as the projection of the standard error ellipsoid on the parameters axes.

1.2 The LIUQE regression

1.2.1 Deriving the Grad-Shafranov equation

The plasma equilibrium in the TCV can be described with the equation of ideal MHD:

$$\begin{aligned}
\boldsymbol{j} \wedge \boldsymbol{B} &= \boldsymbol{\nabla}p \\
\boldsymbol{\nabla} \wedge \boldsymbol{B} &= \mu_0\boldsymbol{j} \\
\boldsymbol{\nabla} \cdot \boldsymbol{B} &= 0
\end{aligned}$$

Combining these equations gives the expression of the magnetic field in cylindrical coordinates:

$$\boldsymbol{B} = -\frac{1}{2\pi r}\frac{\partial\psi}{\partial z}\nabla r + \frac{1}{2\pi r}\frac{\partial\psi}{\partial r}\nabla z + T\nabla\phi \qquad (14)$$

Those equations can be combined in a differential equation similar to the Poisson equation:

$$\Delta^*\psi = -2\pi\mu_0 r j_\phi \qquad (15)$$

with the definition of the operator and the current:

$$\Delta^* = r\frac{\partial}{\partial r}\frac{1}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2} \qquad (16)$$

$$j_\phi = 2\pi\left(\frac{\mathrm{d}p}{\mathrm{d}\psi} + \frac{T}{\mu_0 r}\frac{\mathrm{d}T}{\mathrm{d}\psi}\right) \qquad (17)$$

where p and T are only functions of ψ. The combination of 16 and 17 gives the Grad-Shafranov equation at the core of the algorithm:

$$\Delta^*\psi = -4\pi^2\mu_0 r\left(rp' + \frac{TT'}{\mu_0 r}\right) \qquad (18)$$

This equation is non-linear, and requires specific algorithmic methods.

1.2.2 Solving with iterations and parametrization

The method used to solve equation 18 is the Picard iterations:

$$\Delta^*\psi^{(n+1)} = -2\mu_0 r\left(j_\phi^{(n)}(r,\psi^{(n)}) + j_e\right) \qquad (19)$$

In the framework of equilibrium reconstruction, the boundary between the plasma volume and the vacuum is updated at each iteration based on $\psi^{(n)}$. The functions $p'^{(n)}(\psi^{(n)})$ and $TT'^{(n)}(\psi^{(n)})$ forming the plasma current density are also adjusted at each iteration to best reproduce available experimental measurements.
The second step of the iterative algorithm for solving the inverse equilibrium problem consists in identifying such functions that best reproduce the available measurements. This is performed by first parametrizing p' and TT':

$$p' = g_p(\psi; a_p) \qquad (20)$$

$$TT' = g_T(\psi; a_T) \qquad (21)$$

Where a_p and a_T become the free parameters. In order to simplify the problem, it is possible to restrict it to linear combinations of functions $g_g(\psi)$ that depend on ψ only, the coefficients of the linear combination a_g becoming the free parameters:

$$j_\phi = 2\pi\left(rp' + \frac{TT'}{\mu_0 r}\right) = \sum_g a_g r^{\nu_g} g_g(\psi) \qquad (22)$$

where $\nu_g = 1$ for the terms contributing to p', $\nu_g = -1$ for those contributing to TT', and g_g is set to 0 outside the domain of definition. Thanks to this parametrization, it becomes a linear regression problem, easier to solve.

Finally, a free vertical shift in the flux distribution to stabilise the algorithm is introduced:

$$j_\phi = \sum_g a_g r^{\nu_g} g_g\left(\psi(r, z+\delta z)\right) \qquad (23)$$

Consequently the new free parameters set is $\{a_g, \delta z\}$.

The most frequent choice for the source term base functions g_g are the three polynomials:

$$\nu_1 = 1 \qquad g_1 = (\psi - \psi_0) \qquad \text{for } p' \qquad (24)$$

$$\nu_2 = -1 \qquad g_2 = (\psi - \psi_0) \qquad \text{for } TT' \qquad (25)$$

$$\nu_3 = -1 \qquad g_3 = (\psi - \psi_0)(\psi - \psi_A) \qquad \text{for } TT' \qquad (26)$$

The aim of the study is to analyse the relevance of this choice.

1.2.3 Matrix system for regression

The values of the set of free parameters ont the inner computational grid are stored in a rectangular matrix:

$$T_{yg} = r_y^{\nu_g} g_g\left(\psi(r_y, r_z)\right)\Delta r\Delta z \qquad (27)$$

With this expression, the expected measurements can be written in matrix notation (cf [1], page 18), as resumed here following the notations of the article:

$$\begin{bmatrix}\psi_f\\ B_m\\ I_a\\ I_s\\ I_p\\ \Phi_t\end{bmatrix} = \begin{bmatrix} M_{fa} & M_{fs} & M_{fy}\cdot T_{yg} & \partial_{z_f}M_{fy}\cdot I_y\\ B_{ma} & B_{ms} & B_{my}\cdot T_{yg} & \partial_{z_m}B_{my}\cdot I_y\\ 1_a & 0 & 0 & 0\\ 0 & 1_s & 0 & 0\\ 0 & 0 & T_{pg} & 0\\ 0 & 0 & T_{tg} & 0\end{bmatrix}\cdot\begin{bmatrix}J_a\\ J_s\\ a_g\\ \delta z\end{bmatrix} \qquad (28)$$

where $I_y = j_\phi^{(n-1)}(r_y, z_y)\Delta r\Delta z$; $T_{pg} = \sum_y T_{yg}$; and 1_a is the identity matrix. J_a and J_s are aditionnal free parameters corresponding to uncertainties on coil currents measurements and vessel currents observer.

This system is solved in a least square sense, each equation being given a weight $w_{\dots}$ inversely proportional to the associated measurement error. In order to improve the algorithm, the block structure of the matrix is used:

$$Y_r = \begin{bmatrix} w_f \psi_f \\ w_m B_m \end{bmatrix} \qquad Y_i = \begin{bmatrix} w_p I_p \\ w_t \Phi_t \end{bmatrix} \qquad Y_e = \begin{bmatrix} w_a I_a \\ w_s I_s \end{bmatrix}$$

$$J_e = \begin{bmatrix} J_a \\ J_s \end{bmatrix} \qquad a_j = \begin{bmatrix} a_g \\ \delta_z \end{bmatrix}$$

which gives a more compact expression (weighted version of equation 28):

$$\begin{bmatrix} Y_r \\ Y_e \\ Y_i \end{bmatrix} = \boldsymbol{AdG} \cdot \begin{bmatrix} a_g \\ J_e \\ \delta_z \end{bmatrix} \tag{29}$$

or, written more simply:

$$\boldsymbol{Y} = \boldsymbol{AdG} \cdot \boldsymbol{a_G} \tag{30}$$

The aim of the study is to obtain the vector of regression coefficients $\boldsymbol{a_G}$. Consequently, the algorithm relies on this single linear regression:

$$\boldsymbol{a_G} = (\boldsymbol{AdG})^{-1}\boldsymbol{Y} \tag{31}$$

All the equations from the previous subsection about linear regression analysis are then applicable with $\boldsymbol{AdG} := \boldsymbol{X}$ and $\boldsymbol{a_G} := \boldsymbol{\beta}$.

The complete algorithm loop is summarized on figure 1.

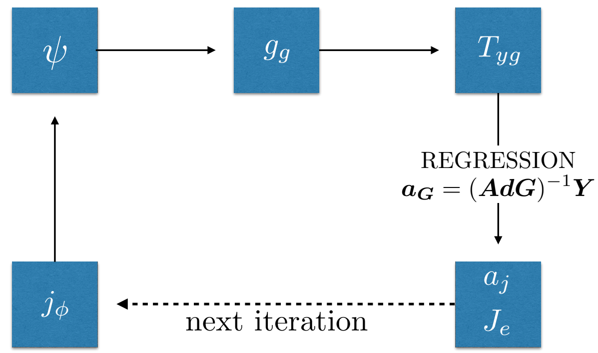

Figure 1: *Algorithm loop.*

2 Results and discussion

The LIUQE algorithm has been used during the whole study exclusively on the TCV shot n° 43760.

2.1 Covariance study for parameters of regression: three basis functions model

2.1.1 Bringing to light the parameters of interest

The covariance matrix (for the model with three basis functions) is computed thanks to equations 6 and 8. The normalization allows to exhibit the correlations without being influenced by the different dimensions of the parameters. The Matlab function `pcolor` is used to show graphically and more intuitively those correlations. The result is given in figure 2.

The main observation is that the parameters 57 to 60 are strongly correlated as the corresponding square submatrix has absolute values near to 1. Those parameters are respectively the three basis function coefficients a_g, and the δz coefficient. The study will consequently focus on it.

It might be useful to note that parameters 1 to 18, which corresponds to coil currents, show some significant correlations between each other aside the diagonal, whereas parameters 19 to 57 corresponding to vessel currents appear to be from far less correlated. More precisely, the eight first parameters (corresponding to E coils measurements) have covariances around 0.2 or 0.3 between each other, whereas parameters 9 to 16 (corresponding to F coils measurements) have stronger correlations but from near to near, not far from the diagonal. Finally, parameters 17 and 18 (corresponding to the big OH coils) are correlated with E coils parameters.

2.1.2 Default of the model with three basis functions

The standard errors computed from the covariance matrix (usual standard error) with the three basis function model are gathered in figure 3a. As described previously, the parameters are grouped by physical interpretation, which allows to give a critical analysis of the regression results.

First, the ratio of standard error on estimates is not good : approximately from 3 to 10% for coil currents (OH-coils excepted), and from 10 to 25% for vessel currents. Note that the fourth parameter has a huge ratio since the value of the estimate β is very small (low current).
The δz_g parameter is also non satisfying with 12% of standard error.
Furthermore, for the parameters of main interest (basis functions), the errors for the three basis function parameters are bigger, especially with the second function which provides a bad 160% of standard error. This shows that a realistic model should use less basis functions, as the algorithm has some difficulties to handle a third function parameter that physically does not exist, and then includes different other contributions in this parameter in order to compute the model.

2.2 Enhancement of the model

As the three basis function model provides unsatisfying results, the LIUQE Algorithm has been applied to the same TCV shot, but without the problematic basis function, i.e. two basis functions, and then with only one. The results for the two cases are given in figures 3b and 3c.

2.2.1 Discussion of the model with two basis functions

First, the ratio of standard error on estimates is not better but in the same order of magnitude: approximately from 4 to 15% for coil currents (OH-coils excepted), and from 10 to 26% for vessel currents. The situation is the same as previously for the fourth parameter.

However, the situation gets better for other aspects of the study. The δz_g parameter has only 5,7% of standard error, which corresponds to a decrease by half by respect to the previous results.
Furthermore, a focus on the basis functions parameters shows clearly that the model has been improved: the standard error is decreased from tenths of percent to around 4% for both parameters. As these ones are the parameters of interests

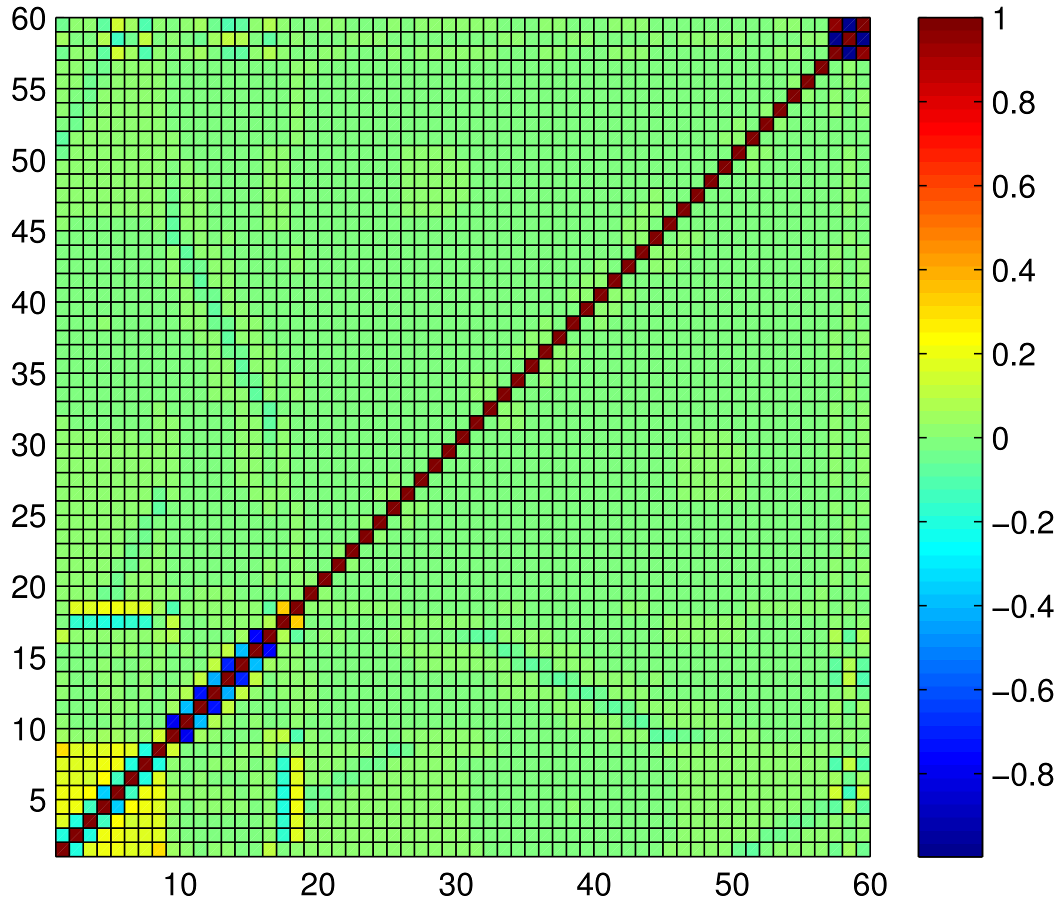

Figure 2: *Normalized covariance matrix $\widehat{\boldsymbol{N}}(\hat{\boldsymbol{\beta}})$ for the β_i parameters with three basis functions, plotted with the `pcolor` Matlab function. Normalized variances are on the diagonal, normalized covariances are on the sides.*

in the study, the model with two basis function seems to be an excellent compromise for the LIUQE algorithm.

2.2.2 Discussion of the model with one basis function

In that case, the ratio of standard error on estimates gets worse as it has strong variations depending on the parameters: approximately from 3 to 44% for coil currents (fourth parameter excluded, as usual), and from 30 to 81% for vessel currents.

The δz_g parameter has 7% of standard error, which is near the result of two basis functions model.
The basis function parameter, as it is the only one remaining in the model, has a standard error about 2,7%.

2.3 Interpretation

Both models are better than the initial one (three basis function) concerning the parameters of interest, i.e. basis functions parameters.

Analyzing these parameters, the model with

#	estimator β	standard error σ from the covariance matrix diagonal	ratio σ/β	maximal standard error from ellipses	
1	-954.57108	119.10218	0.12477036	160.43227	COIL CURRENTS
2	-1875.8882	122.15401	0.065117958	78.606921	
3	-3512.4724	122.01665	0.034738108	76.509117	
4	169.43149	117.49673	0.69347637	486.67099	
5	3428.1035	118.20867	0.034482233	677.91189	
6	3433.743	122.24912	0.035602291	197.9437	
7	835.51655	123.05131	0.14727573	390.65163	
8	-2690.8084	118.30614	0.043966764	88.185406	
9	860.60377	106.23182	0.12343871	182.99931	
10	923.92958	103.63022	0.11216246	232.46261	
11	2013.5353	96.479024	0.047915239	404.15599	
12	1944.6468	96.50297	0.049624934	345.27807	
13	906.32041	96.918641	0.1069364	640.7011	
14	865.22933	97.500362	0.11268731	572.88333	
15	1434.2652	104.74229	0.073028538	162.18759	
16	2028.2714	109.04869	0.053764348	527.55102	
17	15513.26	154.99705	0.0099912622	78.118026	
18	15549.691	152.32989	0.0097963291	79.398221	
19	-1936.1201	175.96136	0.09088349	44.3114	VESSEL CURRENTS
20	-1577.6476	176.12338	0.11163671	47.693486	
21	-1558.2609	176.06217	0.11298632	45.894657	
22	-1531.1164	176.09105	0.11500827	37.92876	
23	-1826.6354	176.20247	0.096462862	49.639238	
24	-1517.0595	176.03153	0.1160347	43.294215	
25	-1199.9641	176.00587	0.14667595	34.417519	
26	-1239.7695	175.82198	0.14181829	69.940709	
27	-1369.1613	175.33201	0.12805796	66.284657	
28	-1205.0379	175.45274	0.14559935	51.372481	
29	-1091.3992	175.50697	0.16080915	62.337221	
30	-982.92793	175.49609	0.17854421	74.660747	
31	-964.71898	175.39567	0.18181012	77.050744	
32	-1080.1413	175.93215	0.16287883	75.513749	
33	-874.25006	175.99953	0.20131487	77.541681	
34	-908.12546	175.76055	0.19354215	62.68214	
35	-842.06431	175.91562	0.20890996	77.568415	
36	-887.6644	175.84988	0.19810401	75.910828	
37	-853.00768	175.91219	0.20622579	78.48909	
38	-855.48186	175.74187	0.20543027	77.280911	
39	-717.62036	175.91259	0.24513322	79.340284	
40	-692.87724	175.84845	0.25379452	76.95047	
41	-687.46233	175.91566	0.25589135	79.273645	
42	-806.03387	175.75701	0.21805164	75.833183	
43	-867.12500	175.99799	0.20296726	77.572992	
44	-1140.3985	175.91918	0.15426115	76.798152	
45	-908.76135	175.38014	0.19298812	77.408385	
46	-990.28444	175.08138	0.17679908	77.037588	
47	-1090.5078	175.06042	0.1605311	75.12861	
48	-1235.4815	175.16186	0.14177619	60.194591	
49	-1391.6539	175.32808	0.12598541	62.465391	
50	-1288.5948	175.83749	0.13645677	67.603636	
51	-1297.4817	176.0046	0.13565093	44.52446	
52	-1343.1809	176.03076	0.13105514	44.674642	
53	-1526.1806	176.19514	0.11544842	45.650401	
54	-1305.3031	176.0899	0.13490346	48.221487	
55	-1444.1261	176.04099	0.1219014	51.871525	
56	-1723.2237	176.12949	0.1022093	61.774278	
57	4070.5897	522.17563	0.12828009	896.01469	BASIS FUNCTIONS
58	794.66598	1276.9832	1.6069434	896.30542	
59	15846.214	9304.5845	0.58718029	896.28106	
60	-0.027377253	0.003392979	0.12393424	896.33235	δz_g

(a) *With three basis functions.*

Figure 3: *Tables of standard errors for each parameter of the linear regression.*

#	estimator β	standard error σ from the covariance matrix diagonal	ratio σ/β	maximal standard error from ellipses	
1	-925.78200	144.57014	0.15616003	86.116666	COIL CURRENTS
2	-1881.086	148.28806	0.078831093	80.354742	
3	-3517.6562	148.21887	0.04213569	80.100816	
4	120.13465	142.8086	1.1887379	80.369467	
5	3316.2542	141.53505	0.042679192	80.370225	
6	3406.0802	148.3487	0.043554083	80.363501	
7	908.61867	148.3115	0.16322744	80.367814	
8	-2648.9489	143.55442	0.054192974	90.446349	
9	970.78118	128.9448	0.13282582	100.99708	
10	864.96144	125.86585	0.14551614	101.00151	
11	1942.7002	116.94679	0.060198064	101.0026	
12	1940.584	116.82492	0.060200912	101.00266	
13	965.99088	117.24459	0.12137236	101.00222	
14	980.61211	118.05021	0.12038421	101.00152	
15	1506.9867	127.43251	0.084561134	100.99583	
16	1879.0617	132.78791	0.070667137	100.97234	
17	15487.078	188.29124	0.012157958	92.174499	
18	15573.265	185.05656	0.011882965	93.334467	
19	-1948.1722	213.75632	0.10972147	76.576072	VESSEL CURRENTS
20	-1575.9069	213.94877	0.13576231	76.584877	
21	-1566.0511	213.85586	0.13655739	76.321907	
22	-1533.0714	213.91326	0.13953248	76.501523	
23	-1821.8946	214.04113	0.11748272	77.698667	
24	-1512.3103	213.84205	0.14140091	76.584327	
25	-1198.9248	213.8108	0.17833546	76.524649	
26	-1236.4336	213.58797	0.1727452	76.637512	
27	-1370.0137	212.99572	0.15546977	80.152218	
28	-1198.5635	213.13816	0.17782802	80.17251	
29	-1083.1675	213.20368	0.19683352	80.098084	
30	-979.30475	213.19142	0.21769671	79.850822	
31	-939.7397	213.06976	0.22673274	76.637751	
32	-1130.5606	213.72304	0.18904165	79.739829	
33	-857.33199	213.80404	0.24938302	76.63774	
34	-891.24762	213.51274	0.23956613	76.637755	
35	-843.35116	213.70179	0.25339598	87.25039	
36	-896.59098	213.6209	0.23825904	80.313954	
37	-853.0212	213.69754	0.25051843	89.947536	
38	-851.71519	213.48947	0.25065829	79.22469	
39	-716.99578	213.6981	0.29804652	92.07948	
40	-693.10807	213.62016	0.30820614	76.63766	
41	-689.55789	213.70168	0.30991116	91.893806	
42	-817.08882	213.50859	0.26130402	76.637752	
43	-882.31332	213.8015	0.24231925	85.516309	
44	-1093.7306	213.70443	0.19539037	79.211913	
45	-937.46125	213.05136	0.22726418	76.637749	
46	-996.12354	212.68869	0.21351637	76.785551	
47	-1096.4651	212.66304	0.19395331	78.964647	
48	-1240.066	212.78603	0.17159251	79.485514	
49	-1393.6168	212.98776	0.15283094	79.693646	
50	-1286.0829	213.60618	0.16609051	79.253395	
51	-1296.3293	213.81	0.16493495	76.609141	
52	-1342.9326	213.84177	0.15923493	76.608881	
53	-1528.0261	214.04077	0.14007665	76.465645	
54	-1306.9948	213.91305	0.16366787	76.49646	
55	-1442.0313	213.85314	0.14829994	52.45015	
56	-1713.2211	213.94523	0.12487893	79.870374	
57	3234.6067	131.69055	0.040713001	94.777575	BASIS FUNCTIONS
58	2384.933	85.486373	0.03584435	95.074259	
59	0.065139077	0.0037476031	0.057532333	101.0044	δz_g

(b) *With two basis functions.*

Figure 3: *Tables of standard errors for each parameter of the linear regression.*

#	estimator β	standard error σ from the covariance matrix diagonal	ratio σ/β	maximal standard error from ellipses	
1	-863.17902	378.14485	0.43808392	249.05641	COIL CURRENTS
2	-1880.1697	387.92532	0.20632463	223.16371	
3	-3542.8321	387.7507	0.10944653	221.79973	
4	64.179295	373.79266	5.8241939	237.72922	
5	3270.755	370.11118	0.11315772	240.14134	
6	3282.8803	388.52883	0.11834998	226.81712	
7	881.15599	387.9159	0.44023522	223.4177	
8	-2539.0794	375.13004	0.14774254	249.43898	
9	1068.1472	337.2772	0.31575909	266.00027	
10	843.25964	329.3011	0.39050974	266.01673	
11	1886.0475	305.9198	0.16220154	266.02075	
12	1740.0573	304.31903	0.17489024	266.02100	
13	682.44938	305.45841	0.44759131	266.01938	
14	795.5444	307.87795	0.38700285	266.01673	
15	1424.5522	332.38565	0.23332642	265.99595	
16	1655.3703	347.92757	0.21018112	265.91036	
17	15438.751	492.51691	0.031901344	250.01944	
18	15625.279	484.06269	0.030979458	250.17711	
19	-1957.3424	559.22428	0.28570591	205.51407	
20	-1568.5356	559.72854	0.35684784	188.87412	VESSEL CURRENTS
21	-1574.6967	559.4926	0.35530181	175.99298	
22	-1550.541	559.6374	0.36093042	187.64145	
23	-1831.2994	559.9711	0.30577801	191.32759	
24	-1507.414	559.45102	0.37113297	200.5148	
25	-1187.3993	559.36779	0.47108651	217.24873	
26	-1233.1311	558.78600	0.45314403	223.27229	
27	-1382.6923	557.25102	0.40301882	223.8501	
28	-1182.8029	557.60014	0.47142269	223.8849	
29	-1068.4215	557.77843	0.52205839	223.91976	
30	-981.2105	557.74992	0.56843044	223.96371	
31	-925.07694	557.43169	0.60257873	233.74272	
32	-1171.0719	559.15392	0.47747188	223.96589	
33	-840.76001	559.35465	0.66529645	245.38596	
34	-897.59834	558.5878	0.62231377	223.9813	
35	-846.14812	559.08403	0.66074014	246.06899	
36	-888.65375	558.87152	0.62889682	223.96554	
37	-853.62071	559.07292	0.65494301	242.57859	
38	-875.95881	558.52044	0.63761039	238.93331	
39	-722.31475	559.07354	0.77400267	234.50444	
40	-689.77376	558.8704	0.81022276	231.29411	
41	-688.41196	559.08369	0.81213535	239.22635	
42	-823.93889	558.57876	0.67793712	223.9895	
43	-892.47297	559.34483	0.62673588	246.20699	
44	-1060.5724	559.09307	0.52716163	228.09837	
45	-958.12386	557.38397	0.58174521	241.95969	
46	-997.06907	556.43373	0.55806939	233.23469	
47	-1097.8748	556.36648	0.5067668	223.96266	
48	-1240.5444	556.6881	0.44874498	223.91297	
49	-1391.8495	557.21586	0.40034203	223.74876	
50	-1280.7429	558.83404	0.43633585	215.77297	
51	-1294.5468	559.36702	0.43209486	210.13876	
52	-1342.3039	559.45018	0.41678355	201.03562	
53	-1530.5621	559.97068	0.3658595	193.69469	
54	-1309.1926	559.63656	0.42746694	194.17694	
55	-1441.8075	559.47991	0.38804063	142.4056	
56	-1706.4714	559.72012	0.32799855	223.52097	
57	5628.2286	152.05219	0.027015994	250.40428	BASIS FUNCTION
58	0.13541078	0.010448707	0.077163037	266.02745	δz_g

(c) *With one basis function.*

Figure 3: *Tables of standard errors for each parameter of the linear regression.*

only one function seems to provide a better accuracy. However, it provides bad results for the currents parameters. The accuracy on the basis function parameter comes from the fact that the algorithm can iterate better and concentrate its accuracy on a single basis function, but only for the corresponding parameter. For the other ones, it lacks one parameter, so the algorithm has to artificially correlate with the other parameters in order to compute the loop iteration.
The three basis function provides non coherent results, especially for one of the basis function, so it is a supernumerary one. Consequently, the algorithm put other contributions from various sources in this parameter (vessel current, etc) in order to complete the regression, what arise new artificial errors.

Consequently, the two basis functions model appears to be more balanced than the one basis function model and the three function model.
This is a coherent result when compared to the Grad-Shafranov equation at the core of the algorithm:

$$j_\phi = 2\pi \left(rp' + \frac{TT'}{\mu_0 r} \right) \tag{32}$$

This current expression involves two free parameters : the density of current and the poloidal component of the current in TCV. It is approximated by a linear combination of the basis functions, so the model with two basis functions fits indeed well the algorithm: one of the functions correspond to the total current (integrated j_ϕ) and another to the TT'.

3 Possible directions for further improvement and research

The analysis of the algorithm regression could be applied on several other TCV shots, in order to verify the validity of the two basis function choice.
The propagation of errors from the current measurements to the basis function coefficient could also be studied more precisely.

References

[1] J.-M. Moret, B.P. Duval, H.B. Le, S. Coda, F. Felici, H. Reimerdes, *Tokamak equilibrium reconstruction code LIUQE and its real time implementation.*

[2] Miyamoto, *Plasma Physics for Nuclear Fusion.*

[3] Sanford Weisberg, *Applied Linear Regression.*

Annexes

Linear regression analysis is a collection of methods whose aim is understanding relations between variables, in a quite simple and very elegant way. The simple regression assumes that two random variables (r.v.) Y and X are linearly connected together with the relationship:

$$Y = \beta_0 + \beta_1 X \tag{33}$$

where β_0 and β_1 are the exact coefficients of the linear equation. One need to add statistical error which can arise from several sources:

$$Y = \beta_0 + \beta_1 X + e \tag{34}$$

For the i-th measurement, the r.v. will take the experimental values y_i, x_i, and e_i:

$$y_i = \beta_0 + \beta_1 x_i + e_i \tag{35}$$

The equation 34 is supposed to represent the *real* physical correlation between Y and X, but it is obviously false. Therefore, the main aim of the experimental data treatment is to estimate the β_j coefficients from a set of measurements, obtaining a simple regression model:

$$Y = \hat{\beta}_0 + \hat{\beta}_1 X + \hat{e} \tag{36}$$

where $\hat{\beta}_0$ and $\hat{\beta}_1$ are the estimates of β_0 and β_1, and $\hat{e}$ is called the residual. The "hat" notation will be used in the following to specify when a mathematical object is an estimate. The method focuses on finding the best estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ minimizing the so called *residual sum of squares* on the whole set of n measurements:

$$RSS(\hat{\beta}_0, \hat{\beta}_1) = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2 = \sum_{i=1}^{n} e_i^2 \tag{37}$$

In the LIUQE case, the regression model has no constant β_0 coefficient, and is extended to multiple variables (p parameters, or r.v.). It still remains linear with respect to all of this new variables:

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + e \tag{38}$$

$$\hat{Y} = \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 + \ldots + \hat{e} \tag{39}$$

These equations are commonly expressed in matrix notation for n measurements and p parameters:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{e} \tag{40}$$

with the vectors and matrix:

$$\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad \boldsymbol{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \tag{41}$$

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1p} \\ x_{21} & x_{22} & \ldots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{np} \end{bmatrix} \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \tag{42}$$

All vectors are written with bold symbols whereas scalar values are written normally. One has to be aware that $\boldsymbol{X}$ is a $n * p$-matrix and $\boldsymbol{\beta}$ a p-vector, whereas $\boldsymbol{Y}$ and $\boldsymbol{e}$ are n-vectors. This notation is very useful for simplification of calculus.